Mathematical Preliminaries to Calculus of Variations

In finite-variable optimization (i.e., ordinary optimization that you most likely know as minimization or maximization of functions), we try to find the extremizing (a term that covers both minimizing and maximizing) values of a finite number of scalar variables to get the extremum of a function that is expressed in terms of those variables. That is, we deal with functions of the form $f(x_1, x_2, \cdots, x_n)$ that need to be extremized by finding the extremizing values of $x_1, x_2, \cdots, x_n$. Calculus of variations also deals with minimization and maximization but what we extremize are not functions but functionals.

The concept of a *functional* is crucial to calculus of variations as is a *function* for ordinary calculus of finite number of scalar variables. The difference between a function and a functional is subtle and yet profound.

Let us first review the notion of a function in ordinary calculus so that we can understand how the functional is different from it.

In this notes, for presenting mathematical formalisms, we will adopt a format that is different from what is usually followed in applied and engineering mathematics books. That is, instead of introducing a number of seemingly unconnected definitions and concepts first and then finally getting to what we really need, here, we will first define or introduce what we actually need and then explain or define the new terms as we encounter them. This takes the suspense out of the notation, definitions, and concepts as they are introduced. New terms are underlined and are immediately explained following their first occurrence. If anything is defined as it is first introduced, it is set in *italics* font.

Because we want to understand the difference between a <u>function</u> and a <u>functional</u>, let us start off with their definitions.

Function

"A rule which assigns a unique real (or complex) number to every $x \in \Omega$ is said to define a real (or complex) *function*."

All is in plain English in the above definition of a function except that we need to say what Ω is. It is called the domain of the function. It is a non-empty open set in $\mathbb{R}^N (\mathbb{C}^N)$.

$\mathbb{R}^N$ (or $\mathbb{C}^N$) is a set of real (or complex) numbers in N dimensions. An element $x \in \mathbb{R}^N$ (or $\mathbb{C}^N$) is denoted by $x = \left\{ x_1, x_2, x_3, \cdots, x_N \right\}$.

While the notion of a set may be familiar to all those who may read this, the notion of an open set may be new to some.

A set S ì $\mathbb{C}^N$ is open if every point (or element) of S is the center of an open ball lying entirely in S.

The open ball with center x_0 and radius r in $\mathbb{R}^N$ is the set $\left\{ x \in \mathbb{R}^N \middle| d_E\left(x_0, x\right) < r \right\}$.

$d_E(x,y) = \left\{ \sum_{i=1}^{n} (x_i - y_i)^2 \right\}^{1/2}$ is the Euclidean distance between $x = \left\{ x_1, x_2, x_3, \cdots, x_N \right\}$ and $y = \left\{ y_1, y_2, y_3, \ldots, y_N \right\}$ both belonging to $\mathbb{R}^N$.

This is how we formally define a function. You can notice how many related concepts are needed to define such a simple thing as a function! One should try to relate to these concepts with one's own prior understanding of what a function is. Let us now do this for a functional so that you can see how it is different so that it too becomes as natural and intuitive as a function is to you. A functional is sometimes loosely (and incorrectly) defined as a function of function(s). But that does not suffice for our purposes because it is subtler than that.

Functional

"A *functional* is a particular case of an <u>operator</u>, in which $R(A)$Î $\mathbb{R}$ or $\mathbb{C}$."
Depending on whether it is real or complex, we define real or complex functionals, respectively.

Are you wondering what $R(A)$ is? Read on to find out.

Operator

A correspondence $A(x)=y$, $x\in X$, $y\in Y$ is called an *operator* from one <u>metric space</u> X into another metric space Y, if to each $x\in X$ there corresponds no more than one $y\in Y$.

The set of all those $x\in X$ for which there exists a correspondence $y\in Y$ is called the *domain* of A and is denoted by $D(A)$; the set of all y arising from $x\in X$ is called the *range* of A and is denoted by $R(A)$.

Thus, $R(A)=\{y\in Y;\ y=A(x),\ x\in X\}$

Note also that $R(A)$ is the *image* of $D(A)$ under the operator A.

Now, what is a metric space?

Metric space

A *metric space* is a pair (X,d) consisting of a set X (of points or elements) together with a metric d, which a real valued function $d(x,y)$ defined for any two points $x,y\in X$ and which satisfies the following four properties:

(i) $d(x,y)\geq 0$ ("non-negative")

(ii) $d(x,y)=0$ if and only if $x=y$ ("zero metric")

(iii) $d(x,y)=d(y,x)$ ("symmetry")

(iv) $d(x,y)\leq d(x,z)+d(z,y)$ where $x,y,z\in X$. ("triangular inequality")

A *metric* is a real valued function $d(x, y), \;\; x, y \in \mathbb{R}^N$ that satisfies the above four properties.

Let us look at some examples of metrics defined in $\mathbb{R}^N$.

1. $d(x,y) = |x - y|$ in $\mathbb{R}$
2. $d(x,y) = \begin{cases} 1 \text{ for } x \neq y \\ 0 \text{ for } x = y \end{cases}$ in $\mathbb{R}$
3. $d(x,y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ in $\mathbb{R}^2$
4. $d(x,y) = |x_1 - x_2| + |y_1 - y_2|$ also in $\mathbb{R}^2$

We can see that the same $\mathbb{R}$ has two different metrics—the first and second ones in the preceding list. Likewise, the third and fourth are two metrics for $\mathbb{R}^2$. Thus, each real number set in N dimensions can have a number of metrics and hence it can give rise to a number of different metric spaces.

The space X we have used so far is good enough for ordinary calculus. But, in calculus of variations, our unknown is a function. So, we need a new set that is made up of functions. Such a thing is called a *function space*. Let us come to it from something more general than that. We call such a thing a vector space. Let us see what this is. First, note that the vector that we refer to here is not limited to what we usually know in analytical geometry and mechanics as something with a magnitude and a direction.

Vector space

A vector space over a field K is a non-empty set X of elements of any kind (called *vectors*) together with two algebraic operations called vector addition ($\oplus$) and scalar multiplication ($\odot$) such that the following 10 properties are true.

1. $x \oplus y \in X$ for all $x, y \in X$. "The set is closed under addition"

2. $x \oplus y = y \oplus x.$ "Commutative law for addition"

3. $(x \oplus y) \oplus z = x \oplus (y \oplus z)$ "Associative law for addition"

4. There exists an additive identity θ such that $x \oplus \theta = \theta \oplus x = x$ for all $x \in X$

5. There exists an additive inverse such that $x \oplus x' = x' \oplus x = \theta$

6. For all $\alpha \in K$, and all $x \in X$, $\alpha \odot \mathrm{x} \in \mathrm{X}$ "The set is closed under scalar multiplication".

7. For all $a \in K$, and all $x, y \in X$, $a \odot (\mathrm{x} \oplus \mathrm{y}) = (a \odot x) \oplus (a \odot y)$

8. $(\alpha + \beta) \odot x = (\alpha \odot x) + (\beta \odot x)$ $\alpha, \beta \in K, \quad x \in X$

9. $(\alpha\beta\ ) \odot x = \alpha \odot (\beta \odot x)$

10. There exists a multiplicative identity such that $1 \odot x = x;$ and $(0 \odot x \in \theta)$

Pardon the strange symbols that are used for addition and multiplication but that generality is needed so that we don't think in terms of our prior notions of usual multiplications and additions. We use the usual symbols to define a *field*, a term we used above.

A set of elements with two binary operators + and • is called a *field* if it satisfies the following ten properties:

1. $a+b=b+a \qquad a,b\in K$
2. $(a+b)+c=a+(b+c) \qquad a,b,c\in K$
3. $a+0=0+a=a \qquad a\in K, \quad (\text{"}0=\text{additive identity"})$
4. $a+(-a)=(-a)+a=0 \qquad (\text{"additive inverse"})$

5. $a \cdot b = b \cdot a$ ("cummutative law")
6. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
7. $a \cdot 1 = 1 \cdot a = a$
8. $a \cdot a^{-1} = a^{-1} \cdot a = 1$ for all $a \in K$ except "0"
9. $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$
10. $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$

Based on the foregoing, we can understand a vector space as a special space of elements (called vectors as already noted) of which the functions that we consider are of just one type.

Next, we consider normed vector spaces, which are simply the counterparts of metric spaces that are defined for normal Euclidean spaces such as $\mathbb{R}^N$.

Normed vector space

A *normed vector space* is a vector space on which a norm is defined.

A *norm* defined on a vector space X is a real-valued function from X to $\mathbb{R}$, i.e., $f : X \to \mathbb{R}$ whose value at $x \in X$ is denoted by $f(x) = \|x\| \in \mathbb{R}$ and has the following properties:

(i) $\|x\| \geq 0$ for all $x \in X$

(ii) $\|x\| = 0$ if and only if $x = \theta$

(iii) $\|\alpha x\| = |\alpha| \|x\|$ $\alpha \in K, \; x \in X$

(iv) $\|x + y\| \leq \|x\| + \|y\|$ $x, y \in X$

The above four properties may look trivial. If you think so, try to think of a norm for a certain vector space that satisfies these four properties. It is not as easy as you may think! Later, we will see some examples of norms for function spaces that we are concerned with in this course.

Let us understand more about function spaces.

Function space

A function space is simply a set of functions. We are interested in specific types of function spaces which are vector spaces. In other words, the "vectors" in such vector spaces are functions. Let us consider a few examples to understand what function spaces really are.

1. $C^0[a,b] \quad a,b \in K; \quad \|x\| = \max_{a \le t \le b} |x(t)|$

As shown above C^0 is a function space of all continuous functions defined over the interval $[a,b]$. It is a normed vector space with the norm defined as shown. Does this norm satisfy the four properties? Please check for yourself.

2. $C_{\text{int}}{}^{0}[a,b] \quad a,b \in K; \quad \|x\| = \int_a^b |x(t)|\,dt$

This represents another function space of all continuous functions over an interval. This too is a normed vector space but with a different norm.

3. $C_{\text{int}\,2}{}^{0}[a,b] \quad a,b \in K; \quad \|x\| = \sqrt{\int_a^b x^2(t)\,dt}$ has yet another norm and denotes one more function space that is a normed vector space.

4. $C^1[a,b] \quad a,b \in K; \quad \|x\| = \max_{a \le t \le b} |x(t)| + \max_{a \le t \le b} |\dot{x}(t)|$

Here, $C^1[a,b]$ is a set of all continuous functions that are also differentiable once. Note how the norm is defined in this case. Does this norm satisfy the four properties? Check it out.

Let us now briefly mention some very important classes of function spaces that are widely used in *functional analysis*—a field of mathematical study of functionals. The functionals are of course our main interest here.